(19) **Europäisches Patentamt / European Patent Office / Office européen des brevets**

(11) **EP 2 357 818 A2**

(12) # EUROPEAN PATENT APPLICATION

(43) Date of publication:
17.08.2011 Bulletin 2011/33

(51) Int Cl.:
H04N 7/15 *(2006.01)*

(21) Application number: **11003903.9**

(22) Date of filing: **08.03.2007**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC MT NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK RS

(30) Priority: **09.03.2006 US 781070 P**
16.03.2006 US 377913

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
07752475.9 / 1 997 313

(71) Applicant: **Citrix Online, LLC**
Santa Barbara, CA 93111 (US)

(72) Inventor: **Christiansen, Bernd**
**Santa Barbara CA 93111 (US)**

(74) Representative: **Hackett, Sean James et al**
Marks & Clerk LLP
Alpha Tower
Suffolk Street
Queensway
Birmingham
B1 1TT (GB)

Remarks:
This application was filed on 12-05-2011 as a divisional application to the application mentioned under INID code 62.

(54) **System and method for dynamically altering videoconference bit rates and layout based on participant activity**

(57) Systems and Methods to transmit and simultaneously display all video streams in a video conference by dynamically adapting video streams based on audio volume of the participants include allocating an amount of bandwidth to a participant in a videoconference; determining a level of activity exhibited by the participant during the videoconference, and changing the amount of bandwidth allocated to the participant responsive to the determined level of activity.

701 — allocating an amount of bandwidth to a participant

703 — determining a level of activity exhibited by the participant

705 — changing the amount of bandwidth allocated to the participant in response to the determined level of activity

Fig. 7

EP 2 357 818 A2
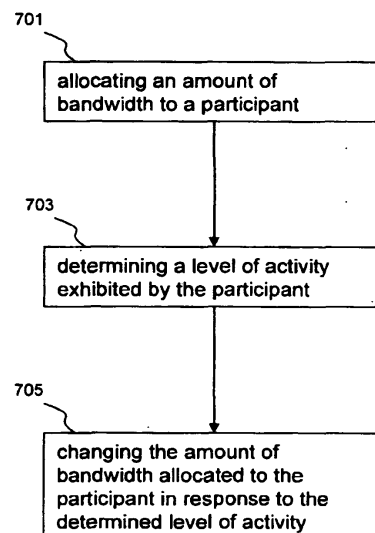
Description

Cross-Reference to Related Application

[0001] The present application claims priority to and the benefit of U.S. Provisional Patent Application No. 60/781,070, filed on March 9, 2006, and U.S. Patent Application No. 11/377,913, filed on March 16, 2006, the disclosures of which are hereby incorporated herein by reference in their entirety.

Field of the Invention

[0002] The present invention relates to video conferences conducted across a network. More particularly, the present invention relates to systems and methods for simultaneously displaying video streams from multiple participants in a video conference and encoding and sending local video input.

Background of the Invention

[0003] Many multi-way video conferencing applications display the video streams of the various participants in a window with a static layout, and may partition the available space on the screen fairly equally among participants. Many layouts are presently used for multi-way video conferencing applications to display the video streams of the various participants in a window. One layout is N-up, which partitions the window into N planar surfaces of equal size. Other video conferencing applications may arrange transformed surfaces in a window. For example, a chat application may feature a three-dimensional view that arranges participants as if they were sitting around a conference-room table.

[0004] In video conferences with many participants, it may become difficult to perceive a speaker as a display screen becomes crowded and views of individual persons become small. Also, many networks have bandwidth limitations which may be exceeded by transmitting a number of video streams above a certain size. Thus, there exists a need for video conference displays that effectively utilize bandwidth and display area to emphasize important information and views.

Summary of the Invention

[0005] The present invention is directed to systems and methods to transmit and simultaneously display all video streams in a large-scale video conference by dynamically adapting video streams based on audio volume of the participants.

[0006] In one aspect, the present invention is a method of allocating bandwidth to participants in a videoconference. The method comprises: allocating an amount of bandwidth to a participant in a videoconference; determining a level of activity exhibited by the participant during the videoconference, and changing the amount of bandwidth allocated to the participant responsive to the determined level of activity.

[0007] In another aspect, the present invention is a method of generating a dynamic graphical representation of a participant in a videoconference that reflects the participant's level of activity during the videoconference. The method comprises: creating a graphical representation of a participant in a videoconference; displaying the created graphical representation; determining a level of activity exhibited by the participant during the videoconference; and altering the displayed graphical representation of the participant responsive to the determined level of activity.

Brief Description of the Drawings

[0008] The foregoing and other objects, aspects, features, and advantages of the invention will become more apparent and may be better understood by referring to the following description taken in conjunction with the accompanying drawings, in which:

[0009] FIGS. 1A and 1B are block diagrams of embodiments of a computing or network device useful as a device in a client-server network;

[0010] FIG. 2 is a block diagram depicting one embodiment of a computer network;

[0011] FIG. 3 is a block diagram depicting one embodiment of a method of generating a dynamic graphical representation of a participant in a videoconference that reflects the participant's level of activity during the videoconference;

[0012] FIG. 4 is a block diagram depicting an embodiment of altering the displayed graphical representation of a videoconference participant responsive to the determined level of activity;

[0013] FIG. 5 is a block diagram depicting another embodiment of altering the displayed graphical representation of a videoconference participant responsive to the determined level of activity;

[0014] FIG. 6 is a block diagram depicting another embodiment of altering the displayed graphical representation of a videoconference participant responsive to the determined level of activity; and

[0015] FIG. 7 is a block diagram depicting one embodiment of a method of allocating bandwidth to participants in a videoconference.

Detailed Description of the Invention

[0016] Figures 1A and 1B depict block diagrams of a typical computer 100 useful as client computing devices and server computing devices. As shown in FIGs. 1A and 1B, each computer 100 includes a central processing unit 102, and a main memory unit 104. Each computer 100 may also include other optional elements, such as one or more input/output devices 130a-130-b (generally referred to using reference numeral 130), and a cache memory 140 in communication with the central process-

ing unit 102.

[0017] The central processing unit 102 is any logic circuitry that responds to and processes instructions fetched from the main memory unit 104. In many embodiments, the central processing unit is provided by a microprocessor unit, such as those manufactured by Intel Corporation of Mountain View, California; those manufactured by Motorola Corporation of Schaumburg, Illinois; the Crusoe and Efficeon lines of processors manufactured by Transmeta Corporation of Santa Clara, California; the lines of processors manufactured by International Business Machines of White Plains, New York; or the lines of processors manufactured by Advanced Micro Devices of Sunnyvale, California.

[0018] Main memory unit 104 may be one or more memory chips capable of storing data and allowing any storage location to be directly accessed by the microprocessor 102, such as Static random access memory (SRAM), Burst SRAM or SynchBurst SRAM (BSRAM), Dynamic random access memory (DRAM), Fast Page Mode DRAM (FPM DRAM), Enhanced DRAM (EDRAM), Extended Data Output RAM (EDO RAM), Extended Data Output DRAM (EDO DRAM), Burst Extended Data Output DRAM (BEDO DRAM), Enhanced DRAM (EDRAM), synchronous DRAM (SDRAM), JEDEC SRAM, PC100 SDRAM, Double Data Rate SDRAM (DDR SDRAM), Enhanced SDRAM (ESDRAM), SyncLink DRAM (SLDRAM), Direct Rambus DRAM (DRDRAM), or Ferroelectric RAM (FRAM). In the embodiment shown in FIG. 1A, the processor 102 communicates with main memory 104 via a system bus 120 (described in more detail below). FIG. 1B depicts an embodiment of a computer system 100 in which the processor communicates directly with main memory 104 via a memory port. For example, in FIG. 1B the main memory 104 may be DRDRAM.

[0019] FIGS. 1A and 1B depict embodiments in which the main processor 102 communicates directly with cache memory 140 via a secondary bus, sometimes referred to as a "backside" bus. In other embodiments, the main processor 102 communicates with cache memory 140 using the system bus 120. Cache memory 140 typically has a faster response time than main memory 104 and is typically provided by SRAM, BSRAM, or EDRAM.

[0020] In the embodiment shown in FIG. 1A, the processor 102 communicates with various I/O devices 130 via a local system bus 120. Various busses may be used to connect the central processing unit 102 to the I/O devices 130, including a VESA VL bus, an ISA bus, an EISA bus, a MicroChannel Architecture (MCA) bus, a PCI bus, a PCI-X bus, a PCI-Express bus, or a NuBus. For embodiments in which the I/O device is an video display, the processor 102 may use an Advanced Graphics Port (AGP) to communicate with the display. FIG. 1B depicts an embodiment of a computer system 100 in which the main processor 102 communicates directly with I/O device 130b via HyperTransport, Rapid I/O, or InfiniBand. FIG. 1B also depicts an embodiment in which local

busses and direct communication are mixed: the processor 102 communicates with I/O device 130a using a local interconnect bus while communicating with I/O device 130b directly.

[0021] A wide variety of I/O devices 130 may be present in the computer system 100. Input devices include keyboards, mice, trackpads, trackballs, cameras, video cameras, microphones, and drawing tablets. Output devices include video displays, speakers, inkjet printers, laser printers, and dye-sublimation printers. An I/O device may also provide mass storage for the computer system 800 such as a hard disk drive, a floppy disk drive for receiving floppy disks such as 3.5-inch, 5.25-inch disks or ZIP disks, a CD-ROM drive, a CD-R/RW drive, a DVD-ROM drive, tape drives of various formats, and USB storage devices such as the USB Flash Drive line of devices manufactured by Twintech Industry, Inc. of Los Alamitos, California.

[0022] In further embodiments, an I/O device 130 may be a bridge between the system bus 120 and an external communication bus, such as a USB bus, an Apple Desktop Bus, an RS-132 serial connection, a SCSI bus, a Fire Wire bus, a Fire Wire 800 bus, an Ethernet bus, an AppleTalk bus, a Gigabit Ethernet bus, an Asynchronous Transfer Mode bus, a HIPPI bus, a Super HIPPI bus, a SerialPlus bus, a SCI/LAMP bus, a FibreChannel bus, or a Serial Attached small computer system interface bus.

[0023] General-purpose computers of the sort depicted in FIG. 1A and FIG. 1B typically operate under the control of operating systems, which control scheduling of tasks and access to system resources. Typical operating systems include: MICROSOFT WINDOWS, manufactured by Microsoft Corp. of Redmond, Washington; MacOS, manufactured by Apple Computer of Cupertino, California; OS/2, manufactured by International Business Machines of Armonk, New York; and Linux, a freely-available operating system distributed by Caldera Corp. of Salt Lake City, Utah, among others.

[0024] For embodiments comprising mobile devices, the device may be a JAVA-enabled cellular telephone, such as the i55sr, i58sr, i85s, or the i88s, all of which are manufactured by Motorola Corp. of Schaumburg, Illinois; the 6035 or the 7135, manufactured by Kyocera of Kyoto, Japan; or the i300 or i330, manufactured by Samsung Electronics Co., Ltd., of Seoul, Korea. In other embodiments comprising mobile devices, a mobile device may be a personal digital assistant (PDA) operating under control of the PalmOS operating system, such as the Tungsten W, die VII, the VIIx, the i705, all of which are manufactured by palmOne, Inc. of Milpitas, California. In further embodiments, the client 113 may be a personal digital assistant (PDA) operating under control of the PocketPC operating system, such as the iPAQ 4155, iPAQ 5555, iPAQ 1945, iPAQ 2215, and iPAQ 4255, all of which manufactured by Hewlett-Packard Corporation of Palo Alto, California; the ViewSonic V36, manufactured by ViewSonic of Walnut, California; or the Toshiba

PocketPC e405, manufactured by Toshiba America, Inc. of New York, New York. In still other embodiments, the mobile device is a combination PDA/telephone device such as the Treo 180, Treo 270, Treo 600, Treo 650, Treo 700, or the Treo 700w, all of which are manufactured by palmOne, Inc. of Milpitas, California. In still further embodiments, the mobile device is a cellular telephone that operates under control of the PocketPC operating system, such as the MPx200, manufactured by Motorola Corp. A typical mobile device may comprise many of the elements described above in FIG. 1A and 1B, including the processor 102 and the main memory 104.

[0025]   Referring now to FIG. 2 one embodiment of a computer network is depicted. In brief overview, a number of clients 213a, 213b, 213c, ... , 213n (collectively referred to as 213), are connected via a network 211, to a each other and to a server 205. The clients 213, networks 211, and server 205 may comprise any computing devices comprising substantially similar capabilities, descriptions, functions, and configurations as described herein.

[0026]   Still referring to FIG. 2, now in greater detail, in the embodiment shown, a number of clients 213 are connected to a network 211. The network 211 may comprise the Internet, local networks, web servers, file servers, routers, databases, computers, servers, network appliances, or any other computing devices capable of sending and receiving information. The network may comprise computing devices connected via cables, IR ports, wireless signals, or any other means of connecting multiple computing devices. The network and any devices connected to the networks may communicate via any communication protocol used to communicate among or within computing devices, including without limitation SSL, HTML, XML, RDP, ICA, FTP, HTTP, TCP, IP, UDP, IPX, SPX, NetBIOS, NetBEUI, SMB, SMTP, Ethernet, ARC-NET, Fiber Distributed Data Interface (FDDI), RS232, IEEE 802.11, IEEE 802.11a, IEE 802.11b, IEEE 802.11g and direct asynchronous connections, or any combination thereof. The network 211 may comprise mobile telephone networks utilizing any protocol or protocols used to communicate among mobile devices, including AMPS, TDMA, CDMA, GSM, GPRS or UMTS.

[0027]   In some embodiments, the network configuration shown may be used to create a videoconference among a plurality of clients 213. In one embodiment, a videoconference may comprise a number of clients 213 transmitting a video stream to a number of other clients 213. In some embodiments, each participant may transmit an a video stream to a server 205, which then retransmits said streams to the other participants. In other embodiments, each participant may transmit a video stream to each other participant individually. In still other embodiments, a combination of centralized and decentralized transmissions may be used. For example, if two participants share a local network, the participants may transmit an audio and video stream directly to each other, and also transmit an audio and video stream to a server

205 for transmission to other non-local participants.

[0028]   A videoconference transmitted over a network may comprise any known or proprietary videoconferencing protocol, including H.320, H.323, H.324, H.225.0, H. 245, H.450, H.235, and H.239 as defined by the International Telecommunications Union, and including any standard or proprietary video compression algorithm, including MPEG-I, MPEG-2, MPEG-4, or H.264. In some embodiments, a client 213 may transmit multiple streams of data as part of a videoconference. For example, a client may transmit one stream of audio data, one stream of video data, one stream of presentation data, and one stream of control data. The audio, video, presentation, and control data may be encoded using any known encoding protocols. In other embodiments, a client may transmit a single stream of data.

[0029]   Referring now to FIG. 3, one embodiment of a method for altering the displayed graphical representation of a participant responsive to the determined level of activity is shown. In brief overview, the method comprises: creating a graphical representation of a participant in a videoconference (step 321); displaying the created graphical representation (step 323); determining a level of activity exhibited by the participant during the videoconference (step 325); and altering the displayed graphical representation of the participant responsive to the determined level of activity (step 327). A videoconference display 300 is also depicted at three time intervals.

[0030]   Still referring to FIG. 3, now in greater detail, the method shown first comprises creating a graphical representation of a participant in a videoconference (step 321). Said graphical representation may be created by any of the computing devices 100 described herein. In one embodiment, the graphical representation of a participant may comprise a graphical display of a video stream received from the participant. A received video stream received may comprise any video stream, and may comprise any known video stream protocol, including MPEG-1, MPEG-2, MPEG-4, H.261, H.263, and M-JPEG.

[0031]   In some embodiments, the graphical representation of a participant may also comprise any display or formatting elements including text, images, borders, and avatars. For example, the video stream of a participant might be displayed along with text indicating the participants name or location. Or, for example, the video stream of a participant might be displayed with an avatar or icon representing the participant.

[0032]   In some embodiments, creating a graphical representation of a participant may comprise transforming the graphical representation such that the graphical representation appears to have three dimensions. In one embodiment, a rectangular video stream may be stretched and skewed to give the appearance of facing a given direction. For example, the video streams of a number of participants may be transformed such that they appear to be arranged around a conference table.

[0033]   After creating a graphical representation of a participant in a videoconference (step 321); the method shown comprises displaying the created graphical representation (step 323). The graphical representation may be displayed on any display device, including a monitor, CRT screen, LCD screen, television, PDA screen, video projector, and laptop screen.

[0034]   In some embodiments, the displaying the created graphical representation may comprise displaying the created graphical representation on a screen also comprising graphical representations of other participants. In these embodiments, any known layout may be used, including N-up, box layouts, ring layouts, and three dimensional layouts. Videoconference screen 300a illustrates on example of an N-up layout with six participants.

[0035]   In some embodiments, the graphical representations of a number of participants may initially be displayed as the same size. In other embodiments, graphical representations of participants may initially be displayed in different sizes. For example, the video stream of a videoconference organizer may be initially allocated a larger screen area.

[0036]   In some embodiments, special screen locations may be assigned to designated participants. For example, the video stream of the videoconference organizer may initially be assigned to the top left position. Or, for example, the video streams of sub groups of participants may be grouped together.

[0037]   In some embodiments, the graphical representations of participants may be displayed on a background. For example, the video stream of a participant might be displayed on a background of a conference table setting, such that the participant appeared to be sitting at the table. r

[0038]   In some embodiments, the participants displayed may all be remote participants. In other embodiments, a videoconference display screen may also comprise a display of the local participant. In some embodiments, the display of the local participant may be specially designated and displayed, according to any of the techniques for altering the graphical representation of a participant described herein.

[0039]   After displaying the created graphical representation (step 323), the method shown comprises determining a level of activity exhibited by the participant during the videoconference (step 325).

[0040]   In some embodiments, the activity level of a participant may be calculated according to the absolute audio volume of the participant. In other embodiments the activity level of a participant may be calculated relative to the audio volumes of all participants. In one embodiment, the activity level of a participant may be computed by dividing the volume output of a given participant by the total volume output from all participants. In other embodiments, other statistical measures, such as deviation from the mean volume of the group, or relative changes in volume from previous levels may be used to compute the activity level of a participant. For example, if a participant's average volume is high but unchanging, this may reflect background noise in the participant's environment rather than participation in the discussion. The activity level of the participant may then be reduced to accommodate this.

[0041]   In some embodiments, said activity level computations may be adjusted to avoid sudden large changes in determined activity level. In one embodiment, the activity level for each participant to each participant is controlled to be within some threshold of its activity level to prevent the activity level from changing too abruptly. In some embodiments, a rolling average of a participant's volume may be used to compute activity level to similarly avoid abrupt changes.

[0042]   In some embodiments, minimum and maximum values for activity level may be set For example, a threshold could be set such that no person accounts for more than 50% of the total of all activity levels, and no person takes accounts for less than 4% of the total of all activity levels.

[0043]   In some embodiments, metrics other than audio volume may be used to determine a participant's activity level. In one embodiment, the activity level of a participant may be computed according to the amount of motion in the participant's video stream. In another embodiment the activity level may be computed according to the amount of motion in the participant's video stream relative to the other participants. In yet another embodiment, the activity level may be computed according to the amount of motion in the participant's mouth area. In still another embodiment, the activity level may be computed according to the amount of motion in the participant's mouth area as compared to other participants. In some embodiments, an activity level may be calculated using both amount of motion and audio volume according to any of the methods described herein.

[0044]   In some embodiments, an activity level may be calculated using information entered by a participant. For, example, a participant momentarily leaving for a cup of coffee may indicate that they are going to be away, and their activity level may be reduced correspondingly. Or, for example, a participant who only intends to watch passively may indicate they are only observing, and their activity level may be reduced accordingly. Or for example, if a user mutes his or her microphone, their activity level may be reduced. In some embodiments, activity levels may be calculated using a combination of user-entered information and any of the other techniques described herein.

[0045]   In other embodiments, activity level may be calculated by any other metric of estimating user participation, including determining the amount of talking or movement of a given participant.

[0046]   In some embodiments, the activity levels may be computed in a centralized manner, such that every participant's display will reflect similar sizes for the participants. In one embodiment, the activity levels may be computed at a server 205 and transmitted to the video-

conference participants. In other embodiments, a client 213 may serve as a master client and compute and transmit activity levels for all participants. In other embodiments, activity levels may be computed in a decentralized manner.

[0047] After determining a level of activity exhibited by the participant during the videoconference (step 325); the method shown comprises altering the displayed graphical representation of the participant responsive to the determined level of activity (step 327). In some embodiments, the displayed graphical representation of a participant may be altered a number of times. In this embodiment, the amount of time between alterations may comprise any time interval, including but not limited to microseconds, milliseconds, seconds, and minutes. In other embodiments, the displayed graphical representation of a participant may be altered continually or continuously over the course of a videoconference.

[0048] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise altering the size of the displayed graphical representation. In some embodiments, the size of the graphical representation of participants exhibiting high activity levels may be increased. In other embodiments, the size of the graphical representation of participants exhibiting low activity levels may be decreased. In some embodiments, each participant may be allocated a percentage of the total screen area based on their percentage of the total activity level. For example, if a participant's activity level accounts for 40% of the total activity level (as measured by adding the activity levels of all participants), that participant may be allocated 40% of the videoconference display area. In some embodiments, a minimum or maximum value may be set for the screen area a given participant may occupy. In some embodiments, the proportion p of the screen corresponding to a given participant may be calculated by using the formula $p = a / \Sigma a$, where a is the activity level of the participant, and Σa is the sum of all the activity levels of all participants.

[0049] Still referring to FIG. 3 a videoconference display 300 in which the displayed graphical representations of the participants are altered over time is shown. A videoconference display 300a is shown at Time 1. The display 300a comprises six areas 310a, 311a, 312a, 313a, 314a, 315a comprising graphical representations of participants in a video conference. In the embodiment shown, each area displays a video stream transmitted by the participant. In the embodiment shown, the areas for the display of video streams are all the same size at the first time point. The time intervals between Times 1, 2 and 3 may be of any magnitude, including but not limited to microseconds, milliseconds, seconds, minutes.

[0050] In the embodiment shown, the displayed graphical representation of the participants has been altered during the time interval between Time 1 and Time 2. For example, the area 315 enlarges, while the areas 311, 312, and 314 are smaller. Said alterations in size may be responsive to activity levels calculated according to

any of the methods described herein. In some embodiments, these changes may reflect that the participant corresponding to area 315 was speaking in the time interval prior to Time 2.

[0051] In the embodiment shown, the displayed graphical representation of the participants has been altered again during the time interval between Time 2 and Time 3. For example, the area 310 enlarges, while the area 313 becomes smaller. Said alterations in size may be responsive to activity levels calculated according to any of the methods described herein. In some embodiments, these changes may reflect that the participant corresponding to area 310 was speaking in the time interval prior to Time 3. In some embodiments, the area 315 maintaining its larger size may reflect that the participant corresponding to area 315 is still actively engaged in conversation.

[0052] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise altering the relative position to other participants of the displayed graphical representation. In one embodiment, participants with high activity levels may be moved higher and closer to the center of a display relative to other participants. In another embodiment, participants with lower activity levels may be moved lower and closer to the edges of a display relative to other participants. For example, the participants may all be initially displayed in a row. Displays of participants may then be raised or lowered according to their activity level. In other embodiments, a designated position may be set aside for participants with activity levels above a given threshold. For example, the participant with the highest activity level may be displayed in the center of a ring, or in the top-left corner of a grid. In other embodiments, designated positions may be set aside for participants with activity levels below a given threshold. For example, participants with an activity level indicating they have not yet spoken may be displayed in a bottom row.

[0053] Referring now to FIG. 4, a block diagram depicting an embodiment of altering the displayed graphical representation of a videoconference participant responsive to the determined level of activity is shown. In brief overview, a videoconference display screen 400, is depicted at three time points, Times 1, 2 and 3. The videoconference display 400 comprises a number of areas 410, 411, 412, 413, 414, 415 in which video streams are displayed. The time intervals between Times 1, 2 and 3 may be of any magnitude, including but not limited to microseconds, milliseconds, seconds, minutes.

[0054] Still referring to FIG. 4, now in greater detail, a videoconference display 400a is shown at a first time point. The display 400a comprises six areas 410, 411, 412, 413, 414, 415 for displaying video streams. Each area corresponds to a participant in the videoconference, and may comprise any of the graphical representations described herein.

[0055] In the embodiment shown, several of the areas 410, 411, 412, 413, 414, 415 and their corresponding

video streams change size and position during the time interval between Time 1 and Time 2. For example, the areas 410, 41 1, 413, and 414 shrink and move downward, while the area 415 enlarges and moves up and to the center of the display 400. Said changes in size and position may be responsive to activity levels calculated according to any of the methods described herein.

[0056] In the embodiment shown, several of the areas 410, 411, 412, 413, 414, 415 and their corresponding video streams again change size and position during the time interval between Time 2, and Time 3. For example, area 415 enlarges further and moves to the top-center of the display, while area 412 shrinks and moves to the bottom-right corner. Said changes in size and position may be responsive to an activity level calculated according to any of the methods described herein.

[0057] Now referring to FIG. 5, a block diagram depicting one embodiment of dynamically computing videoconference display areas is shown. In brief overview, a videoconference display 500 is depicted at 2 time points, Time 1 and Time 2. The display 500 is divided into a plurality of circles 510, 511, 512, 513, wherein each circle corresponds to a videoconference participant. The time intervals between Times 1 and 2 may be of any magnitude, including but not limited to milliseconds, seconds, minutes, and hours.

[0058] Still referring to FIG. 5, now in greater detail, each circle, such as 510 corresponds to a given videoconference participant. In the embodiment shown, at Time 1, all circles are the same size and have the same radius. In the embodiment shown, the video stream of each participant would be displayed inside the circle corresponding to that participant. In some embodiments, a participant's video stream may be displayed in an area comprising a rectangle inscribed in said participant's circle.

[0059] In the embodiment shown the size of each circle may be computed by treating the circle as an enclosed cylinder holding gas at a given pressure. The pressure may correspond to the activity level of the given participant, and may be calculated according to any of the methods described herein. Once the pressure is computed, the radius of the circle can be computed using known formulae based on physical principles. In one embodiment, the radius, r, of a circle may be computed using the formula

$$r = \sqrt{(kp/\pi)}$$

where p is the pressure of the circle, and k is an initial constant. In some embodiments, the initial constant k may be set such that all circles are initially the same size and together occupy the entire display. In other embodiments, circles may be set with different initial constants to provide for different initial sizes. Other formulae may also be used to compute the subsequent position of the circles.

[0060] In the embodiment shown, the circles 510, 511, 512, 513 change size between Time 1 and Time 2. For example, the circle 511 becomes smaller, while the circle 513 grows larger. Said changes in pressure and corresponding size and position may be responsive to activity levels calculated according to any of the methods described herein.

[0061] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise altering the transformation of the displayed graphical representation. In one embodiment, transformations of the displayed graphical representation may be altered such that the representations appear to be facing the participant with the highest activity level.

[0062] Now referring to FIG. 6, a block diagram depicting another embodiment of altering the displayed graphical representation of a videoconference participant responsive to the determined level of activity is shown. In brief overview, a videoconference display screen 600, is depicted at three time points, Times 1, 2 and 3. The videoconference display 600 comprises a number of areas 610, 611, 612, 613 in which video streams are displayed. The time intervals between Times 1, 2 and 3 may be of any magnitude, including but not limited to milliseconds, seconds, minutes, and hours.

[0063] Still referring to FIG. 6, now in greater detail, a videoconference display 600a is shown at a first time point. The display 600a comprises four areas 610, 611, 612, 613 corresponding to participants in a videoconference. Each area may comprise any of the graphical representations of participants discussed herein. In the embodiment shown the rectangular areas are transformed such that they appear to be part of a three-dimensional scene.

[0064] At Time 1, in the embodiment shown, the areas 610, 611, 612, 613 are displayed in an arch centered around the area 612. In the embodiment shown, the area 612a is transformed so that it and its corresponding video stream appear flat. In the embodiment shown, the areas 610a, 611 a, and 613a are transformed such that the areas appear to be facing area 612a. Said transformations may be performed according to any known image and video transformation algorithms, and may be implemented in accordance with any technique described herein.

[0065] At Time 2, in the embodiment shown the areas 610, 611, 612, 613 have changed transformation. For example, area 613 is transformed so that it appears at a greater angle, while area 610 is transformed so that it appears to be more directly facing outward. Said changes may be responsive to activity levels calculated according to any of the methods described herein. In some embodiments, said changes may be a result of more active participation by the participant corresponding to area 610.

[0066] At Time 3, in the embodiment shown, the areas 610, 611, 612, 613 have again changed their transformation. For example, area 610 is transformed so that it

appears to be directly facing outward, while areas 611, 612, and 613 are transformed such that they appear to be facing area 610. Said changes may be responsive to activity levels calculated according to any of the methods described herein.

[0067] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise altering the graphical representation of a participant in other ways. In one embodiment, participants below a given activity level threshold may be displayed in black and white, while persons above a given threshold are displayed in color. In another embodiment, participants below a given threshold may be displayed in grayscale. In some embodiments, the bit-depth of the colors used in the graphical representation of a participant may be altered. In other embodiments, a colored border may be used to indicate the participant with the highest activity level, or the participant most likely to be speaking. In other embodiments, any attributes of a participant's video stream may be altered, including but not limited to color, brightness, contrast, and sharpness. For example, the video stream of a relatively inactive participant may be dimmed and/or blurred, while the video stream of an active participant may be brightened and/or sharpened.

[0068] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise displaying a still image in place of a video stream for a given participant. In some embodiments, participants below a given activity level threshold may be displayed as a static image. In some embodiments, said static image may represent the frame of the participant's video stream prior to their activity level falling below the threshold. In other embodiments, the static image may comprise a still photo or avatar designated by the participant.

[0069] In some embodiments, altering the displayed graphical representation of a participant (step 327) may comprise altering text within the graphical representation. In some embodiments, a participant with an activity level below a given threshold might be displayed with text such as "idle" or "inactive". In other embodiments, a participant with an activity level above a given threshold might be displayed with text such as "active" or "speaking".

[0070] In any of the embodiment described above for altering the displayed graphical representation of a participant, methods may be used so that the displayed graphical representation of a participant does not change too rapidly. For example, the altered relative position of a participant may be limited to a certain distance from the previous position. Or for example, the altered transformation of a participant may be limited to a certain angle from the previous transformation.

[0071] Referring now to FIG. 7, a method of allocating bandwidth to participants in a videoconference is shown. In brief overview, the method comprises:

[0072] allocating an amount of bandwidth to a participant in a videoconference (step 701); determining a level of activity exhibited by the participant during the videoconference (step 703); and changing the amount of bandwidth allocated to the participant responsive to the determined level of activity (step 705).

[0073] Still referring to FIG. 7, now in greater detail, the method shown comprises allocating an amount of bandwidth to a participant in a videoconference (step 701). Said allocation may be done by any allocation method. In some embodiments, bandwidth may be initially allocated equally to all participants. In other embodiments, designated participants, such as the videoconference organizer, may be initially allocated greater bandwidth than other participants. In still other embodiments, bandwidth may be allocated based on local connection bandwidth. For example, if a total bandwidth for a videoconference is fixed, bandwidth may be allocated to individual participants based on the capabilities of the participants local connections.

[0074] After allocating an amount of bandwidth to a participant in a videoconference (step 701), the method comprises determining a level of activity exhibited by the participant during the videoconference (step 703). This level of activity may be determined using any of the methods described herein.

[0075] After determining a level of activity exhibited by the participant during the videoconference (step 703), the method comprises changing the amount of bandwidth allocated to the participant responsive to the determined level of activity (step 705). Any known means of changing video transmission bandwidth may be used. In some embodiments, the bandwidth may be changed by transmitting at a lower or higher resolution. In one embodiment, the resolution of the transmitted video may correspond to the current screen area allocated to the participant. In other embodiments, the bandwidth may be changed by transmitting at a lower or higher frame rate. In still other embodiments, the bandwidth may be changed by transmitting video with more or fewer colors. In still other embodiments, the bandwidth may be changed by transmitting video in black and white or grayscale. In yet other embodiments, the bandwidth may be changed by altering the quantization of prediction errors or the granularity of the subsampling. In one embodiment, changing the bandwidth may comprise switching to a different video codec.

[0076] In one embodiment, changing the amount of bandwidth allocated to the participant responsive to the determined level of activity (step 705) may comprise each participant may adjusting the bit rate of his or her own video stream such that the video stream consumes the same fraction of the globally available bandwidth as the ratio of the participant's screen estate to the total screen estate. The globally available bandwidth may either be static, assigned by a server 205 at run-time, or collaboratively estimated at run-time using well-known techniques. In another embodiment, a participant may change the bandwidth of his or her own video stream in response to the participants activity level as a percentage

of the total activity level. In other embodiments, a user's local bandwidth may also be included in calculating the bit rate of a given video stream.

[0077] In one embodiment, if a participant's activity is below a given threshold, the participant's video-stream bit-rate may be reduced to zero. In this embodiment, a still image may be displayed in place of the participant's video stream.

[0078] In some embodiments, the techniques described herein may be employed in videoconferences wherein each participant transmits a local video stream to all other participants. In one embodiment, each participant may utilize any of the above methods in a decentralized manner. For example, a participant may receive a number of video streams from other participants, and compute the activity level and corresponding size of the displayed representation of each participant using any of the methods described above. Or, for example, a participant may receive a number of video streams from other participants, compute the activity level of each participant, compute the proportion the participant's own activity level to the total, and adjust the participant's own transmitted bandwidth accordingly using any of the methods described herein.

[0079] In other embodiments, the techniques described herein may be employed in videoconferences wherein any of the discussed methods are performed in a centralized manner. For example, a server 205, or elected master client 213 may compute activity levels for each participant and transmit said activity levels to all other participants. Or, for example, a server 205 may compute both activity levels and corresponding bandwidth allocations for each participant, and transmit information corresponding to said bandwidth allocations to all participants. In some embodiments, the server 205 may then alter the bandwidths that the server 205 retransmits the video streams of said participants.

[0080] In some embodiments, an interface may provide for manual override of any of the functions described herein. In some embodiments, a given participant may enlarge the display of a given participant if the user so wishes. For example, in some embodiments, functionality may be provided so that a participant may request that the user's video stream is displayed at a given size.

[0081] While the invention has been particularly shown and described with reference to specific preferred embodiments, it should be understood by those skilled in the art that various changes in form and detail may be made therein departing from the spirit and scope of the invention as defined by the appended claims and in the paragraphs below.

In one aspect related to the invention, there is disclosed a method of generating a dynamic graphical representation of a participant in a videoconference that reflects the participant's level of activity during the videoconference. The method comprises the steps of: (a) creating a graphical representation of a participant in a videoconference; (b) displaying the created graphical representation; (c)

determining a level of activity exhibited by the participant during the videoconference; and (d) altering the displayed graphical representation of the participant responsive to the determined level of activity.

In a method embodying this aspect, step (a) may comprise one or more of the following operations: creating a two-dimensional graphical representation of a participant in a videoconference; creating a three-dimensional graphical representation of a participant in a videoconference; creating a two-dimensional graphical representation of a participant in a videoconference using an image of the participant; and/or creating a two-dimensional video representation of a participant in a videoconference.

In a method embodying this aspect, step (c) may comprise one or more of the following operations: determining the participant's absolute volume level during the videoconference; determining the participant's volume level during the videoconference relative to other participants; determining the participant's volubility during the videoconference; determining the participant's level of animation during the videoconference; and/or determining the participant's level of animation during the videoconference relative to other participants.

In a method embodying this aspect, step (d) may comprise one or more of the following operations: changing the size of the displayed graphical representation of the participant responsive to the determined level of activity; changing the relative position of the displayed graphical representation of the participant responsive to the determined level of activity; changing the color bit depth of the displayed graphical representation of the participant responsive to the determined level of activity; replacing a color graphical representation of the participant with a grayscale representation of the participant responsive to the determined level of activity; replacing a color graphical representation of the participant with a black and white representation of the participant responsive to the determined level of activity; replacing a video representation of the participant with a still image responsive to the determined level of activity; changing the z-axis ordering of the displayed graphical representation of the participant responsive to the determined level of activity; and/or changing the transformation of the displayed graphical representation of the participant responsive to the determined level of activity.

In a second aspect related to the invention, there is disclosed an apparatus for generating a dynamic graphical representation of a participant in a videoconference that reflects the activity level of the participant during the videoconference. The apparatus comprises: a display engine displaying a graphical representation of a participant in a videoconference; an activity monitor in communication with the display engine, the activity monitor determining a level of activity exhibited by the participant during the videoconference, wherein the display engine alters the displayed graphical representation of the participant responsive to the determined level of activity.

In an apparatus embodying the second aspect, the display engine may create one or more of the following: a two-dimensional graphical representation of a participant in a videoconference; a three-dimensional graphical representation of a participant in a videoconference; a two-dimensional graphical representation of a participant in a videoconference using an image of the participant; and/or a two-dimensional video representation of a participant in a videoconference.

In an apparatus embodying the second aspect, the the activity monitor may determine one or more of the following: the participant's absolute volume level during the videoconference; the participant's volume level during the videoconference relative to other participants; the participant's volubility during the videoconference; the participant's level of animation during the videoconference; and/or the participant's level of animation during the videoconference relative to other participants.

In an apparatus embodying the second aspect, the display engine may perform one or more of the following operations: change the size of the displayed graphical representation of the participant responsive to the determined level of activity; change the relative position of the displayed graphical representation of the participant responsive to the determined level of activity; change the color bit depth of the displayed graphical representation of the participant responsive to the determined level of activity; replace a color graphical representation of the. participant with a grayscale representation of the participant responsive to the determined level of activity; replace a color graphical representation of the participant with a black and white representation of the participant responsive to the determined level of activity; replace a video representation of the participant with a still image responsive to the determined level of activity; change the z-axis ordering of the displayed graphical representation of the participant responsive to the determined level of activity, and/or change the transformation of the displayed graphical representation of the participant responsive to the determined level of activity.

In a third aspect related to the invention, there is disclosed a method of allocating bandwidth to participants in a videoconference. The method comprises the steps of:

(a) allocating an amount of bandwidth to a participant in a videoconference;
(b) determining a level of activity exhibited by the participant during the videoconference; and (c) changing the amount of bandwidth allocated to the participant responsive to the determined level of activity.

In an method embodying the third aspect, step (b) may comprise one or more of the following operations: determining the participant's absolute volume level during the videoconference; determining the participant's volume level during the videoconference relative to other participants; determining the participant's volubility during the

videoconference; determining the participant's level of animation during the videoconference; and/or determining the participant's level of animation during the videoconference relative to other participants.

In a method embodying the third aspect, step (b) may comprise the operation of (b-a) receiving a communication from another participant in the video conference, the communication including an indication of the activity level exhibited by the another participant; and (b-b) comparing an indication of the participant's level of activity to the another participant's level of activity.

In a method embodying the third aspect, step (c) may comprise one or more of the following operations: increasing the amount of bandwidth allocated to the participant; and/or decreasing the amount of bandwidth allocated to the participant. Further, said method may comprise one or more of the following operations: changing the size of the transmitted graphical representation of the participant responsive to the bandwidth allocated to the participant; changing the color bit depth of the transmitted graphical representation of the participant responsive to the bandwidth allocated to the participant; replacing a color graphical representation of the participant with a grayscale representation of the participant responsive to the bandwidth allocated to the participant; replacing a color graphical representation of the participant with a black and white representation of the participant responsive to the bandwidth allocated to the participant; and/or replacing a video representation of the participant with a still image responsive to the bandwidth allocated to the participant.

In a fourth aspect related to the invention, there is disclosed an apparatus for allocating bandwidth to participants in a videoconference. The apparatus comprises: a network engine transmitting data to a remote participant in a videoconference; an activity monitor in communication with the network engine, the activity monitor determining a level of activity exhibited by a local participant during the videoconference, wherein the network engine changes the available bandwidth responsive to the determined level of activity.

In an apparatus embodying the fourth aspect, the activity monitor may determine one or more of the following: the local participant's absolute volume level during the videoconference; the local participant's volume level during the videoconference relative to other participants; the local participant's volubility during the videoconference; the local participant's level of animation during the videoconference; and/or the local participant's level of animation during the videoconference relative to other participants.

The apparatus embodying the fourth aspect may further comprise a display engine in communication with the network engine, wherein the display engine may be capable of performing one or more of the following operations: changing the size of the transmitted graphical representation of the participant responsive to the amount of bandwidth allocated to the local participant; changing the color

bit depth of the transmitted graphical representation of the participant responsive to the amount of bandwidth allocated to the local participant; replacing a color graphical representation of the participant with a grayscale representation of the participant responsive to the amount of bandwidth allocated to the local participant; replacing a color graphical representation of the participant with a black and white representation of the participant responsive to the amount of bandwidth allocated to the local participant; and/or replacing a video representation of the participant with a still image responsive to the amount of bandwidth allocated to the local participant.

Claims

1. A method of allocating bandwidth to a participant in a videoconference, the method comprising:

    (a) allocating an amount of bandwidth to a participant in a videoconference;
    (b) determining a level of activity exhibited by the participant during the videoconference; and
    (c) changing the amount of bandwidth allocated to the participant responsive to the determined level of activity.

2. The method of claim 1 wherein step (b) comprises (b-a) receiving a communication from another participant in the videoconference, the communication including an indication of the activity level exhibited by the another participant; and (b-b) comparing an indication of the participant's level of activity to the another participant's level of activity.

3. The method of claim 1 wherein step (c) comprises increasing tH amount of bandwidth allocated to the participant.

4. The method of claim 1 wherein step (c) comprises decreasing the amount of bandwidth allocated to the participant.

5. An apparatus allocating bandwidth to participants in a videoconference, the apparatus comprising:

    a network engine transmitting data to a remote participant in a videoconference;
    an activity monitor in communication with the network engine, the activity monitor determining a level of activity exhibited by a local participant during the videoconference;

    wherein the network engine changes the available bandwidth responsive to the determined level of activity.

6. The apparatus of claim 5 wherein the activity monitor
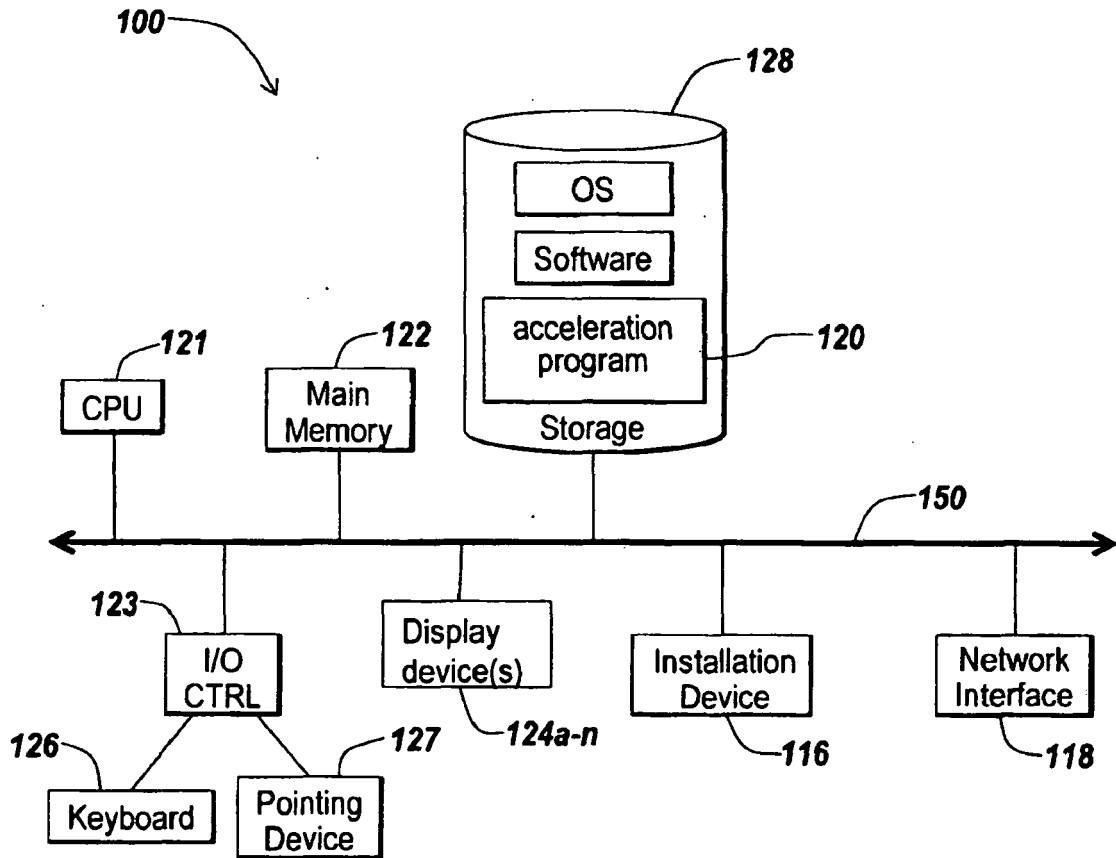
determines any one of:

    (a) the local participant's absolute volume level during the videoconference;
    (b) the local participant's volume level during the videoconference relative to other participants;
    (c) the local participant's volubility during the videoconference;
    (d) the local participant's level of animation during the videoconference; or
    (e) the local participant's level of animation during the videoconference relative to other participants.

7. The apparatus of claim 5 further comprising a display engine in communication with the network engine, the display engine either changing the size of or changing the color bit depth of the transmitted graphical representation of the participant responsive to the amount of bandwidth allocated to the local participant.
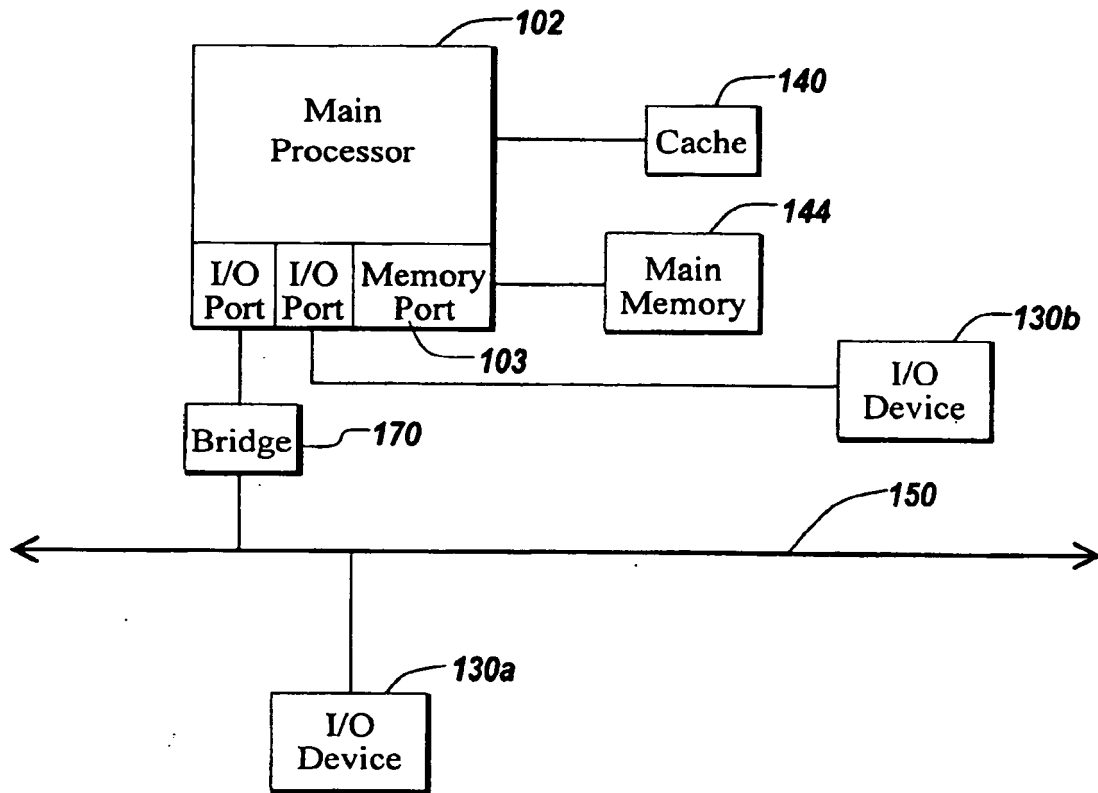
8. The apparatus of claim 5 further comprising a display engine in communication with the network engine, the display engine replacing a color graphical representation of the participant with either a grayscale or a black and white representation of the participant responsive to the amount of bandwidth allocated to the local participant.

9. The apparatus of claim 5 further comprising a display engine in communication with the network engine, the display engine replacing a video representation of the participant with a still image responsive to the amount of bandwidth allocated to the local participant.
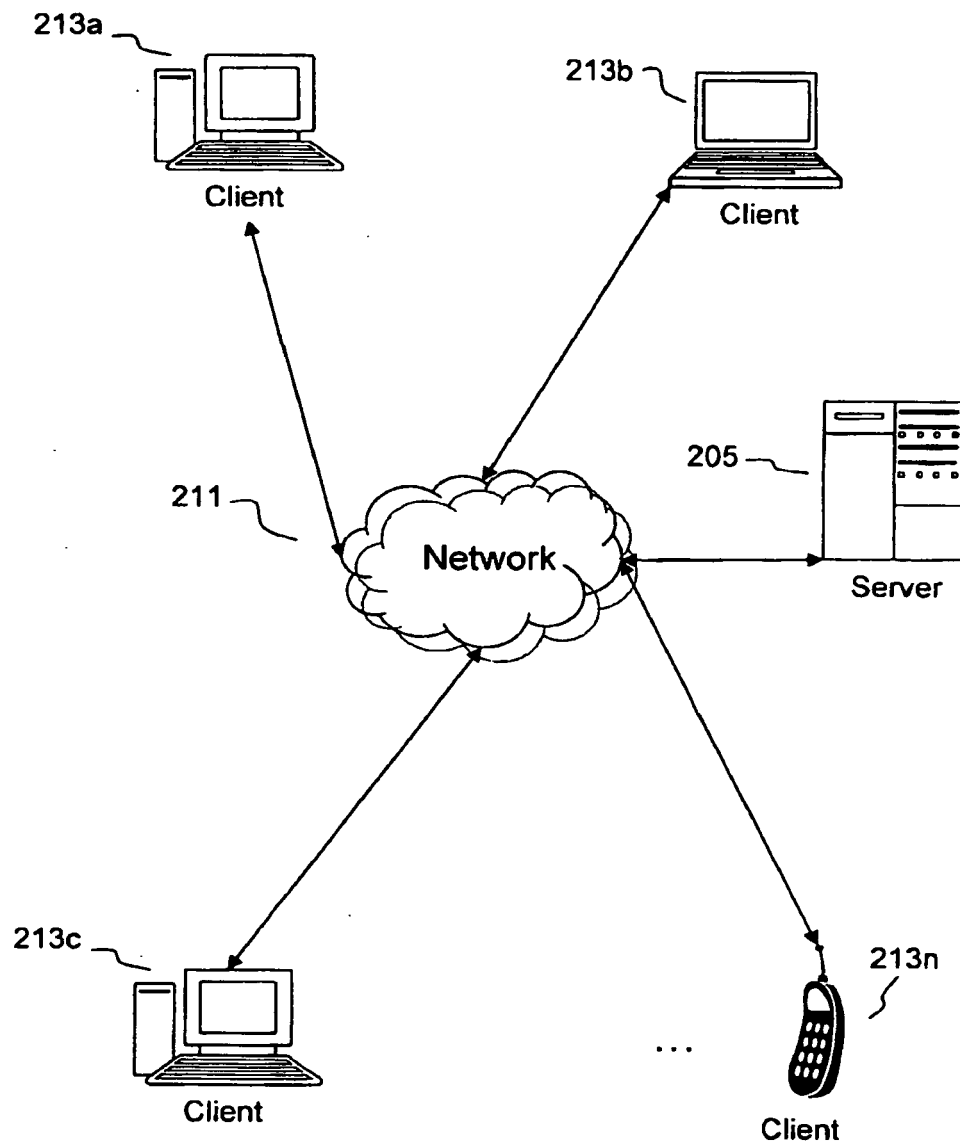
10. The method of claim 1 further comprising implementation of an apparatus as defined in any of claims 5-9.
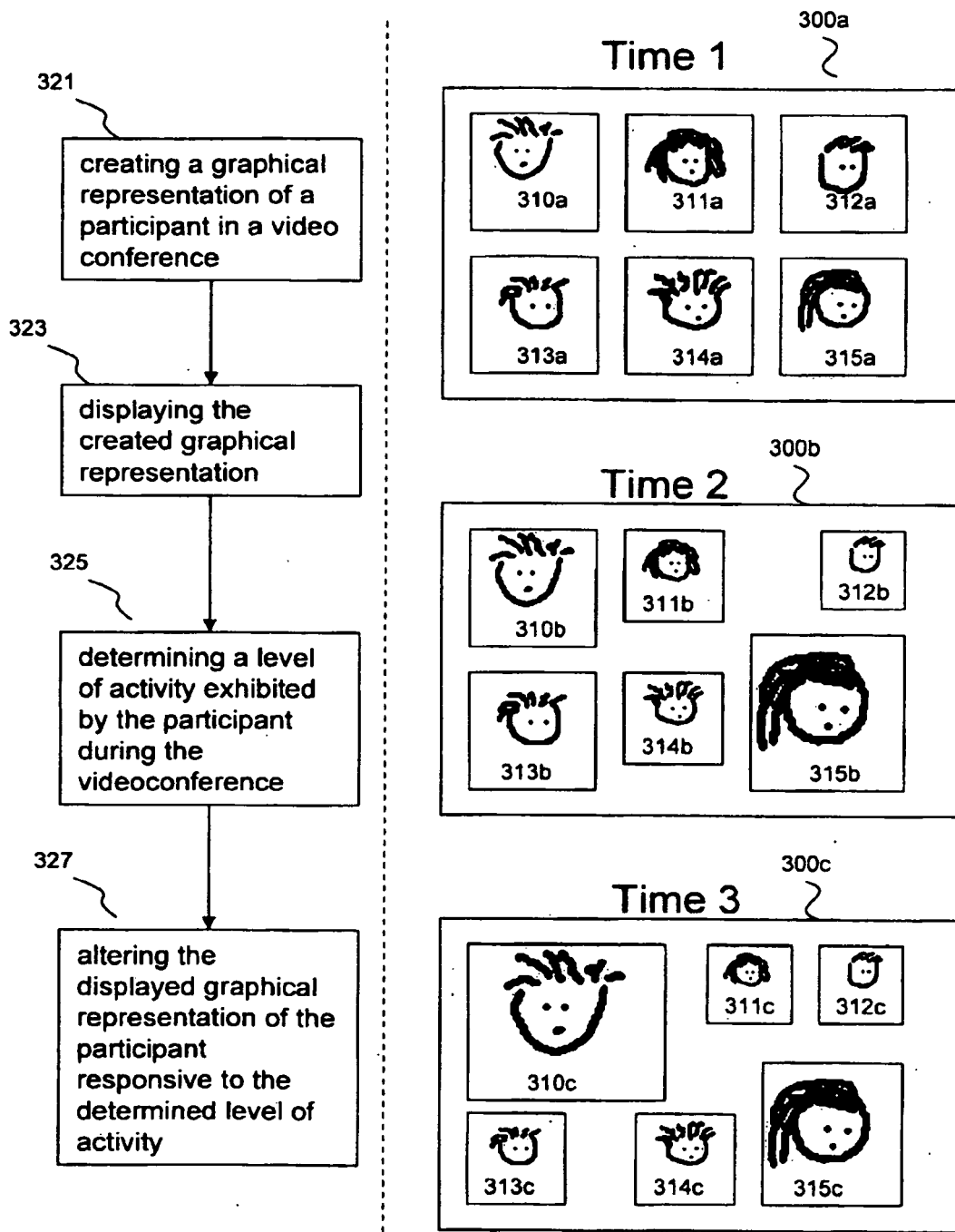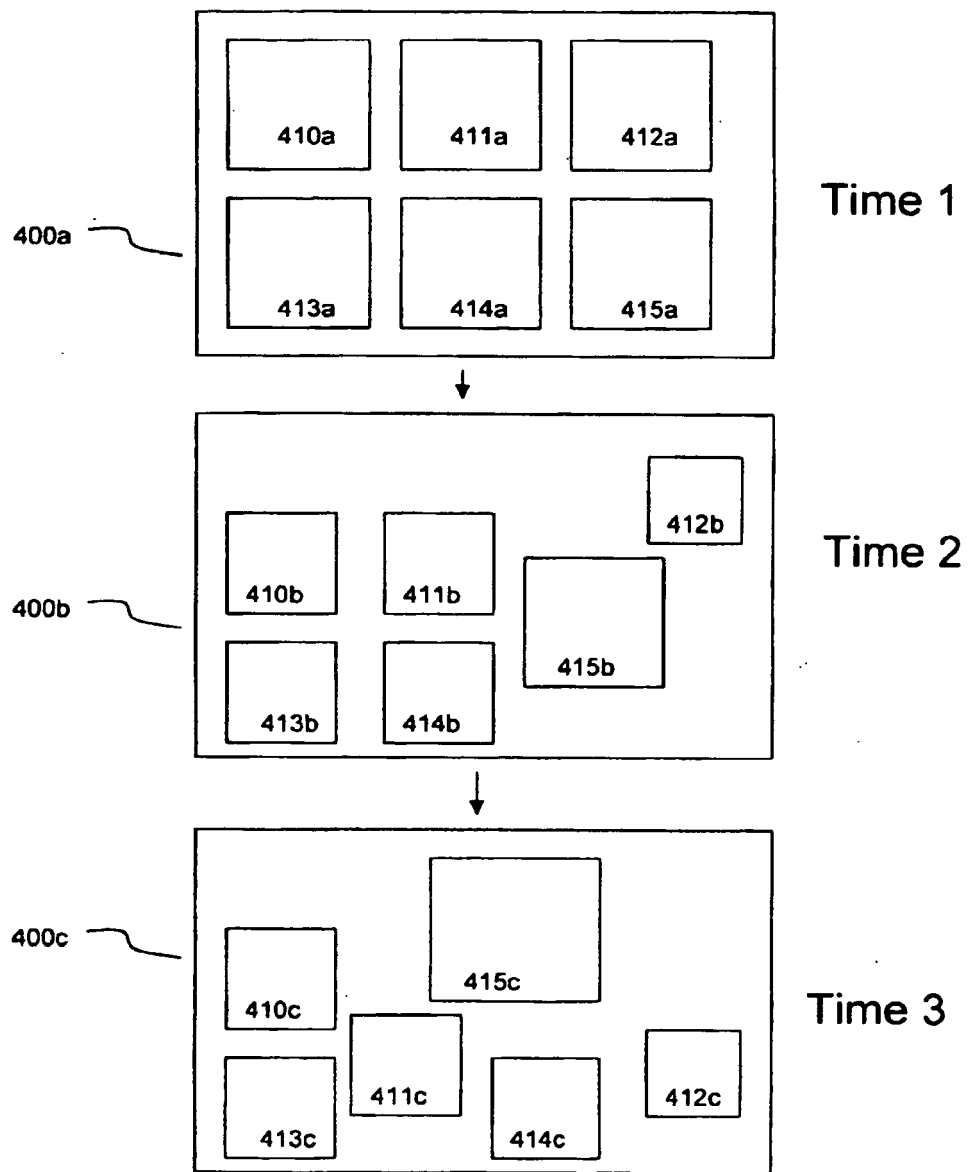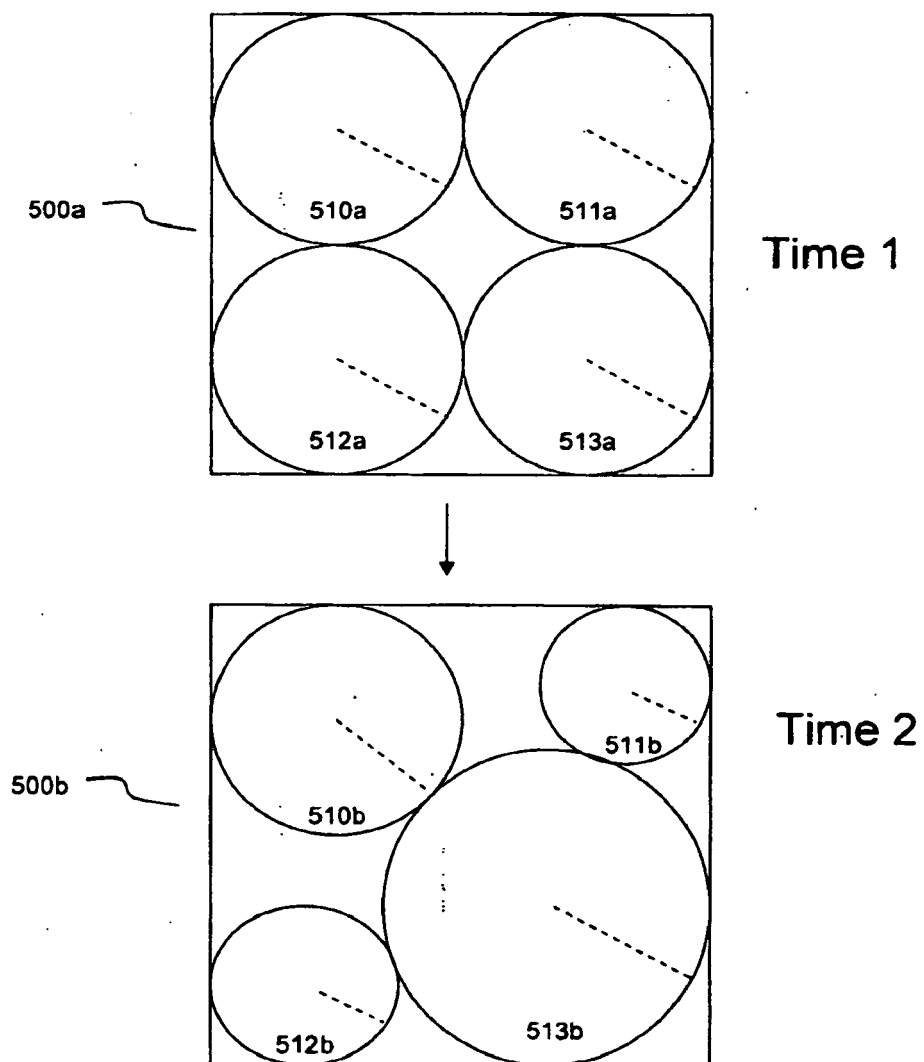
Fig. 1A

Fig. 1B

213a

213b

211

205

Network

Server

213c

. . .

213n

Client

Client

Client

Client

Fig. 2

321

creating a graphical
representation of a
participant in a video
conference

323

displaying the
created graphical
representation

325

determining a level
of activity exhibited
by the participant
during the
videoconference

327

altering the
displayed graphical
representation of the
participant
responsive to the
determined level of
activity

Time 1   300a

310a   311a   312a

313a   314a   315a

Time 2   300b

310b   311b   312b

313b   314b   315b

Time 3   300c

310c   311c   312c

313c   314c   315c

Fig. 3

400a

410a 411a 412a

413a 414a 415a

Time 1

400b

410b 411b 412b

413b 414b 415b

Time 2

400c

410c 415c

413c 411c 414c 412c

Time 3

Fig. 4

Fig. 5

600a

Time 1

610a 611a 612a 613a

600b

Time 2

610b 611b 612b 613b

600c

Time 3

610c 611a 612a 613a

Fig. 6

701

allocating an amount of
bandwidth to a participant

703

determining a level of activity
exhibited by the participant

705

changing the amount of
bandwidth allocated to the
participant in response to the
determined level of activity
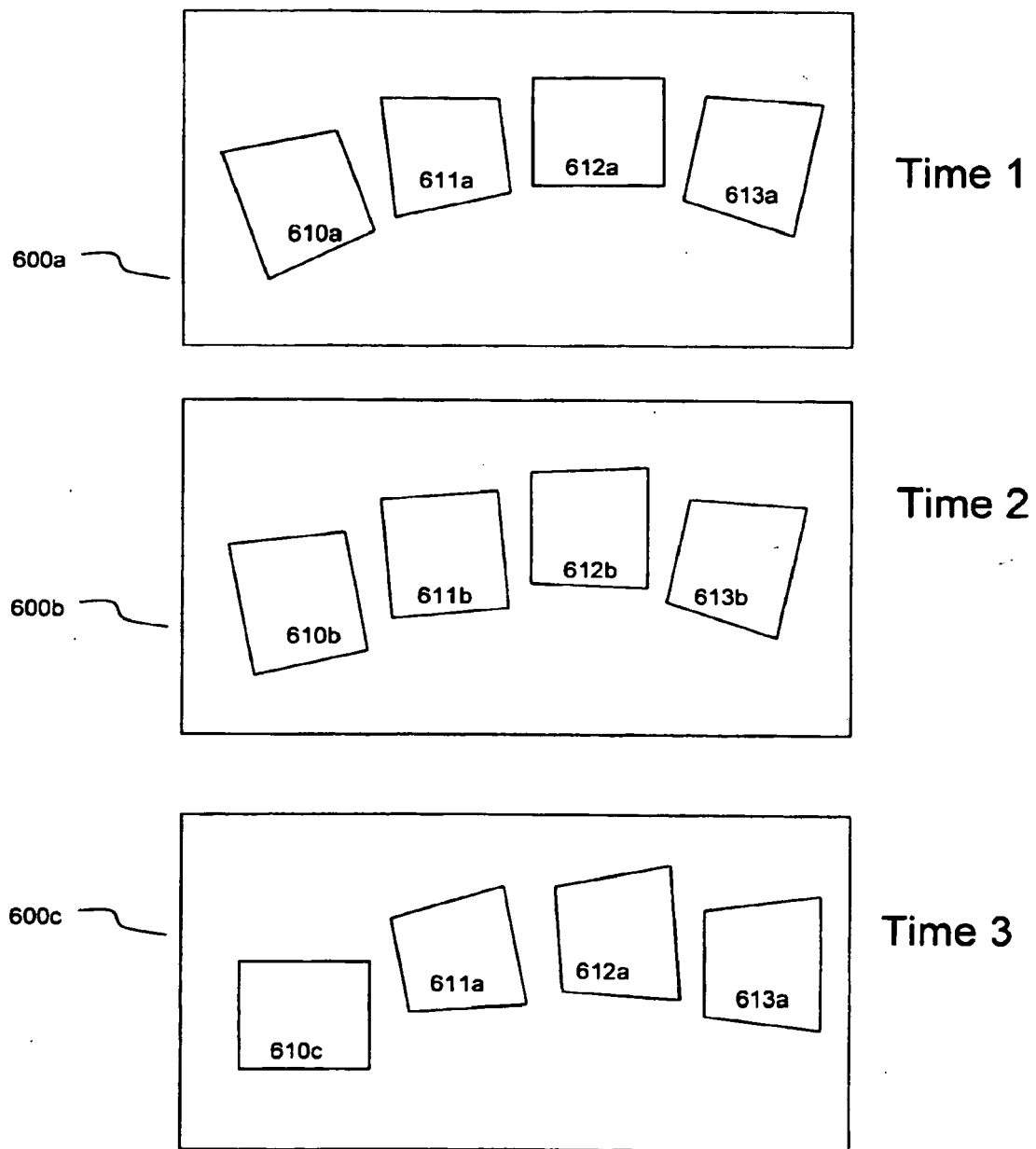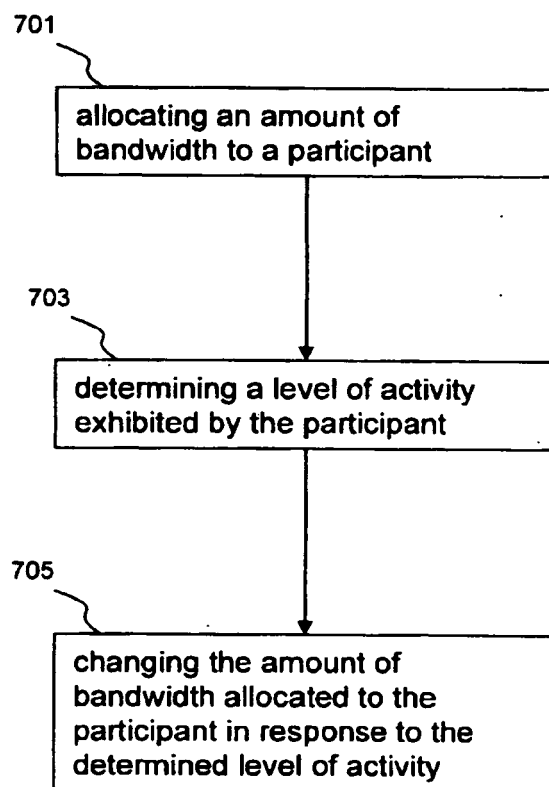
Fig. 7

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 78107006 P **[0001]**
- US 37791306 A **[0001]**